MULTIMEDIA UNIVERSITY

STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 2, 2018/2019

DSP5018 – STATISTICS AND PROBABILITY

(*for diploma students only*)

12 MARCH 2019

2.30 pm – 4.30 pm

(2 Hours)

INSTRUCTIONS TO STUDENT

1. This question paper consists of **3** pages excluding cover page and appendix.

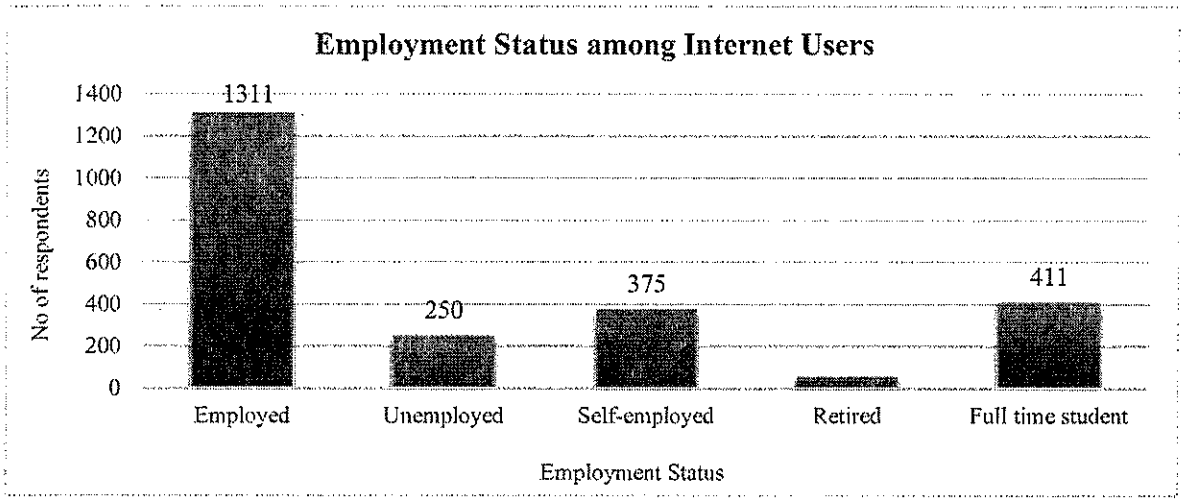
2. Attempt **ALL FOUR (4)** questions. The marks distribution for each question is given.

3. Write your answers in the answer booklet provided.

4. Key formulae are given in the Appendix.

Question 1

a) The following graph shows the distribution of internet users by employment status.
Source: Malaysian Communications & Multimedia Commission (MCMC) - Internet Users Survey 2017.



i. From a survey of 2402 internet users, find the number of retired respondents. (1 mark)

ii. Present the above information in a table which include the classes, frequencies, relative frequencies and percentage. (6 marks)

b) The following stem-and-leaf diagram shows the amount spent (in RM) for fuel expenses in a monthly basis of 50 students in a private college.

Stem	Leaf
10	1 1 2 5 6 9
20	0 2 3 3 4 6 7 8 9 9 9
30	0 1 1 2 3 4 5 6 7 7 8 9 9
40	1 2 4 5 6 6 7 8
50	3 5 6 6 9
60	2 3 4 5
70	2
80	7 9
Key: 10	1 = 101

i. Determine the median and mode of the above data. (3 marks)

ii. Identify the type of variable in the above situation. (1 mark)

iii. Based on the stem-and-leaf diagram above, construct a grouped-data-table where the first class starts at $100 \leq 200$. (6 marks)

iv. From table in part (iii), compute the sample mean and standard deviation. Give your answer in two decimal places. (8 marks)

[TOTAL 25 MARKS]

Continued...

Question 2

a) A survey has been carried out among 900 undergraduates in Multimedia University. The survey result showed that 575 undergraduates owed a smartphone and 225 undergraduates owned a tablet. However, there are 675 undergraduates owned a smartphone or a tablet. If an undergraduate is chosen at random, what is the probability that

i. this undergraduate owned a smartphone and a tablet? (3 marks)

ii. this undergraduate owned a tablet, given that he also owned a smartphone. (4 marks)

b) The probability of a referee making a serious mistake in a match is 0.21. Assuming that the event is independent, what is the probability that the referee

i. does not make any serious mistake in all the five matches? (4 marks)

ii. makes at least one serious mistake in three matches? (4 marks)

c) In a particular day, the probability that it will be a raining day is 0.32. If it rains, Susan will drive her own car to the workplace, otherwise she will take a bus to the workplace. If she drives her own car, the probability that she will be late to the workplace is 0.24. If she takes the bus, the probability that she will be late to the workplace is 0.13.

i. Construct a tree diagram to represent the above information. (3 marks)

ii. What is the probability that she will be late to the workplace? (3 marks)

iii. If Susan late to the workplace, what is the probability that it is a raining day? (4 marks)

[TOTAL 25 MARKS]

Question 3

a) A basketball coach claimed that the average height of his players is greater than 170cm. A sample of 37 players were selected and their average height is 172.24cm with a standard deviation of 7cm. Based on the sampled data, can we accept the coach's claim at 1% significance level? (10 marks)

b) A private college in Melaka has found that the proportion of students who accept the offer to further study at the college is 0.7. After a new marketing head is appointed, the college wants to check whether the proportion has changed significantly. If 876 out of 1200 students accepted the offer, can it be concluded at 5% significance level, that there has been a real change for the acceptance rate? (12 marks)

c) Calculate the p-value for a one tail test with test statistic $z_0 = 1.47$. (3 marks)

[TOTAL 25 MARKS]

Continued...

Question 4

Kaels Company would like to study the relationship between advertising expenses and sales revenue for the first eight months of 2017. The following data is recorded.

Advertising Expenses (RM million), x	3	10	7	6	5	7	12	10
Sales Revenue (RM million), y	25	35	32	30	27	31	42	38

a) Given that $\sum x^2 = 512$, $\sum y^2 = 8672$, $\sum xy = 2065$. Calculate the value of SS_{xx} , SS_{yy} and SS_{xy}. (3 marks)

b) Calculate the correlation coefficient, r and interpret the answer. (3 marks)

c) Find the linear regression equation $\hat{y} = \beta_0 + \beta_1 x$. *(Round to **4 decimal** places in all your calculations and final answer)* (4 marks)

d) Predict the sales revenue if the advertising expenses is RM 30 million. (2 marks)

e) Compute the value of *SST*, *SSR* and *SSE*. (3 marks)

f) Calculate the standard error of the estimate, s_e . (2 marks)

g) At 5% significance level, do the data provide sufficient evidence to conclude that advertising expenses and sales revenue are positively linearly correlated? (8 marks)

[TOTAL 25 MARKS]

End of page.

APPENDIX – KEY FORMULA

	Ungrouped Data	Grouped Data
Mean	$\bar{x} = \frac{\sum x_i}{n}$ where n : sample size	$\bar{x} = \frac{\sum m_i f_i}{\sum f_i}$ where m : class midpoint, f : class frequency
Variance	$s^2 = \frac{1}{n-1}\left[\sum x_i^{\,2} - \frac{\left(\sum x_i\right)^2}{n}\right]$ where n : sample size	$s^2 = \frac{1}{\left(\sum f_i\right)-1}\left[\sum m_i^{\,2} f_i - \frac{\left(\sum m_i f_i\right)^2}{\sum f_i}\right]$ where m : class midpoint, f : class frequency

NORMAL AND STANDARD NORMAL PROBABILITY DISTRIBUTION

- z-value (observed value) for an x value : $Z = \frac{x-\mu}{\sigma}$

HYPOTHESIS TESTING

Population	Mean	Proportion
1	$Z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$	$Z = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$ where $\hat{p} = \frac{x}{n}$
2	$Z = \frac{\left(\bar{x}_1-\bar{x}_2\right)-\left(\mu_1-\mu_2\right)}{\sqrt{\frac{\sigma_1^{\,2}}{n_1}+\frac{\sigma_2^{\,2}}{n_2}}}$	$Z = \frac{\left(\hat{p}_1-\hat{p}_2\right)-\left(p_1-p_2\right)}{\sqrt{\hat{p}_p\left(1-\hat{p}_p\right)}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ where $\hat{p}_p = \frac{x_1+x_2}{n_1+n_2}$

Continued...

LINEAR REGRESSION ANALYSIS

- **Least square regression equation,** $\hat{y} = \beta_0 + \beta_1 x$ where $\beta_1 = \dfrac{S_{xy}}{S_{xx}}$ and $\beta_0 = \bar{y} - \beta_1 \bar{x}$

- **Correlation coefficient,** $r = \dfrac{S_{xy}}{\sqrt{S_{xx} S_{yy}}}$ where

$S_{xx} = \sum x^2 - \dfrac{(\sum x)^2}{n}$	$S_{yy} = \sum y^2 - \dfrac{(\sum y)^2}{n}$	$S_{xy} = \sum xy - \dfrac{(\sum x)(\sum y)}{n}$

- **Sum of Square**

$SST = SSR + SSE$	$SST = S_{yy}$	$SSR = \dfrac{(S_{xy})^2}{S_{xx}}$

- **Test of Significance for Regression Slope**

Standard Error, $s_e = \sqrt{\dfrac{SSE}{n-2}}$	Regression t-Test $t = \dfrac{\beta_1}{s_e / \sqrt{Sxx}}$	Correlation t-Test $t = r \cdot \sqrt{\dfrac{n-2}{1-r^2}}$

End of page.